Construction and Validation of Parentage Testing for Thoroughbred Horses by 53 Single Nucleotide Polymorphisms

Kei-ichi HIROTA¹⁾, Hironaga KAKOI²⁾, Hitoshi GAWAHARA²⁾, Telhisa HASEGAWA³⁾ and Teruaki TOZAKI^{1)*}

Departments of ¹⁾Molecular Genetics and ²⁾Genetic Analysis, Laboratory of Racing Chemistry, 1731-2 Tsurutamachi, Utsunomiya, Tochigi 320-0851 and ³⁾Equine Research Institute, Japan Racing Association, 321-4 Tokami-cho, Utsunomiya, Tochigi 320-0856, Japan

(Received 29 October 2009/Accepted 24 January 2010/Published online in J-STAGE 3 February 2010)

ABSTRACT. We characterized the SNP 53 JPN System for parentage verification during horse registry. The SNP 53 JPN System was constructed using 53 highly polymorphic single nucleotide polymorphisms (SNPs), which were amplified and genotyped with 2 multiplex assays. The SNP 53 JPN System showed good resolution for 95 unrelated thoroughbreds, and the exclusion probability (PE01) for each SNP ranged from 11.5 to 23.0%, resulting in a total PE01 value of 99.996%. These results indicate that the SNP 53 JPN System is useful for parentage testing of thoroughbreds. Of the 53 SNPs, 8 SNPs could be used to exclude a pseudo parent and sib combination found using the 2006 International Society for Animal Genetics (ISAG) horse comparison test, as efficiently as the parentage testing systems using short tandem repeats (STRs). Thus, we concluded that the SNP 53 JPN System could provide sufficient and reliable information for routine parentage testing of thoroughbred.

KEY WORDS: horse, microsatellite, short tandem repeat, single nucleotide polymorphism, thoroughbred.

J. Vet. Med. Sci. 72(6): 719–726, 2010

Horse breed registries worldwide rely on genetic testing to ensure the reliability of studbooks. At present, parentage testing of horses is carried out using short tandem repeats (STRs) [19, 34, 35], also known as microsatellites, with the StockMarks® for Horses Equine 17-plex Genotyping Kit (Applied Biosystems, Foster City, CA, U.S.A.) and the Equine Genotypes™ Panel 1.1 (Finnzymes Diagnostics, Espoo, Finland) as basic parentage systems, and the Microsatellite 15 TKY System [33] as an additional parentage testing system for resolving single-marker exclusion issues. Because these systems can be used to identify a large number of alleles, the total exclusion probability (PE01) of each system is above 99.99%, which fulfills the requirement (>99.95%) of the International Stud Book Committee (ISBC) for horse parentage testing. Although the use of STRs sometimes results in slippage mutation errors, these genotyping systems are abundantly used by horse registries worldwide.

Single nucleotide polymorphisms (SNPs) are biallelic sequences abundantly dispersed throughout most eukaryotic genomes [23, 26, 29]: they have the following characteristics: (i) they have very low mutation rates compared to STRs (10^{-8} vs. 10^{-3}); (ii) because of their short length they are highly suitable for analysis using automated high-throughput technologies; (iii) they can be detected by several techniques, such as denaturing high-performance liquid chromatography (DHPLC) [20], invader assays [22], and TaqMan® assays [12]. Therefore, there is growing interest in SNPs in the field of forensics [17, 18].

Recently, there has been increased interest in the use of SNPs for parentage testing of domestic animals [28, 36],

and with the completion of the horse genome project, numerous SNPs have also become available for genetic studies of horses. However, there is insufficient information on their utility, such as their discrimination power and the ability to be combined in multiplex assays.

In this study, we designed 2 multiplex assay panels: JPN_01 (27 SNPs) and JPN_02 (26 SNPs). The assay panels were constructed for genotyping a total of 53 SNPs using the Sequenom MassARRAY® system [14] by the single-base primer extension method. The system was designed to maintain the PE01 above 99.95%, which is a requirement of the ISBC for horse parentage testing systems. Then, we evaluated the efficacy of the 2 genotyping systems—one utilizing STRs and the other utilizing SNPs.

MATERIALS AND METHODS

Animals: Genomic DNA was isolated from whole blood samples collected from 93 unrelated thoroughbreds in Japan. It was extracted using the MagExtractor system MFX-2000 (Toyobo, Osaka, Japan) according to the manufacturer's protocol. Two DNA samples, HCT06_01 and HCT06_05, from a pseudo parent and a sib, respectively, obtained from the 2006 International Society for Animal Genetics (ISAG) horse comparison test, were used to evaluate the exclusion probabilities.

Assay design for the genotyping of 53 SNPs: One hundred and twenty SNPs—representing all horse chromosomes—were selected from the horse SNP collection, EquCab2.0 (<http://www.broad.mit.edu/mammals/horse/>), and genotyped by direct sequencing for 8 thoroughbreds in order to evaluate their allele frequencies. Of the 120 SNPs, 60 SNPs that showed higher heterozygosities were used for designing the multiplex assays. The iPLEX assay was designed using the MassARRAY® Assay Design software (Sequenom Inc.,

* CORRESPONDENCE TO: TOZAKI, T., Department of Molecular Genetics, Laboratory of Racing Chemistry, 1731-2 Tsurutamachi, Utsunomiya, Tochigi 320-0851, Japan.
e-mail: ttozaki@lrc.or.jp, ttozaki@nyc.odn.ne.jp

San Francisco, CA, U.S.A.), and 2 multiplex assay panels of 27 and 26 SNPs each were finally constructed to genotype 53 SNPs in total (Table 1).

Microsatellite genotyping: To compare the discrimination powers of SNPs and STRs in horse parentage testing, 2 STR genotyping systems were used in this study. The first system contains 15 STRs from the marker set reported by Kakoi *et al.* [15], which includes the minimum 9 markers recommended by the ISAG, and is used for routine parentage testing; the *LEX3* and *TKY321* markers [15] were excluded because *LEX3* is located on a sex chromosome and

TKY321 is also contained in the Microsatellite 15 TKY System. The second system used was the Microsatellite 15 TKY System [33], which is used in many laboratories as an additional parentage testing system to resolve single-marker exclusion issues. Alleles are designed with alphabetical symbols, in the order of smallest to largest, based on a middle-sized allele having been assigned as M. Those definitions have been confirmed by ISAG horse comparison test.

Statistical analysis: Allele frequencies of each SNP and STR were determined by direct counting from tests performed for 95 thoroughbreds. Heterozygosity (H) was cal-

Table 1. Primer and extend primer list for the SNP 53 JPN system

SNP name	Panel	PCR primer (forward)	PCR primer (reverse)	Extend primer
BIEC607490	JPN_01	ACGTTGGATGTTGACACTGAACGCATTCCC	ACGTTGGATGTTCCAGAAGCTTCTGTACCCC	GCATTCCTCCCTCA
BIEC397614	JPN_01	ACGTTGGATGCATGTCTGGTAGACGTGTCC	ACGTTGGATGGGAAAATATAGACATGTCTCT	cACGTGTCCCAAGAAC
BIEC555737	JPN_01	ACGTTGGATGAGCTTATTTCTGGCTTCTG	ACGTTGGATGACAGCTCTCCAGCAAAATC	TGGCTTCTGTCTTTT
BIEC871916	JPN_01	ACGTTGGATGAGGAGAGGAAGACAGATGTG	ACGTTGGATGAACCGCATCTTTGTGGTGTG	ATGTGACAGAGGGAGTT
BIEC870244	JPN_01	ACGTTGGATGTGAAGTCTCTCAGTCTAC	ACGTTGGATGCTTAGGAAAGCTGTGGACTC	CCATGCAGAAAAATCCAG
BIEC696480	JPN_01	ACGTTGGATGTTTCACTACCAAGCAACAG	ACGTTGGATGATTTATCACCACGGCCAGTC	cccAGCCAGACACATCC
BIEC581695	JPN_01	ACGTTGGATGATGTTTCCACCAATCGTGAG	ACGTTGGATGTAAGAAACCTAGAGTGAAC	CAGAAGCCCAATTTGAAT
BIEC672139	JPN_01	ACGTTGGATGACTTCAGTTGTCTGGTTTGTG	ACGTTGGATGTCATCTCCCTCTAGCTATG	GGTTTGTGACACAATGATG
BIEC322734	JPN_01	ACGTTGGATGATGAGCAGTGAAAATCTGGGC	ACGTTGGATGAATTAGGCAGCCTCAGTGTG	cgaATCTAATTCAGACCCC
BIEC681989	JPN_01	ACGTTGGATGCCTCATATTTCCCAAGACCC	ACGTTGGATGGGTGGGAGTTTGGAGATTAC	CCAAGACCCATTTCAAATAG
BIEC532060	JPN_01	ACGTTGGATGAGCAGGATCAATCAGACAG	ACGTTGGATGCTGCACGTTTTCTCTGTGAC	gggtAGACAGGAGCCACTAC
BIEC604433	JPN_01	ACGTTGGATGGTCTACGCATATTAAGAGGG	ACGTTGGATGGAATAAGCTCAGGTCTTTTG	gcTCTTCAGAATGCCAAAGAT
BIEC911841	JPN_01	ACGTTGGATGTCTTGACATGACAGCCATCG	ACGTTGGATGGCTGCTCTCTTGTCTTCTG	cctaTGCTTCCATTATAGCCC
BIEC205561	JPN_01	ACGTTGGATGATCAGGTAAAGTCCAGGTG	ACGTTGGATGGGAAAAGTGAATCCAGGAAAC	gTCCATTATGCACCTGAACCTT
BIEC508410	JPN_01	ACGTTGGATGATCTACTGATCCTGAGACTG	ACGTTGGATGGAGAAGCTGTAAGTTTGGG	gggtATCTCTGAGACTGCTGTG
BIEC633181	JPN_01	ACGTTGGATGCCAAAGGCTAAGAGCAAGAG	ACGTTGGATGGCAAGATGAAAGTTCGTCC	ccCCTAAAGAACCTGACTGAAAA
BIEC344848	JPN_01	ACGTTGGATGTCTACTCCACAGGGTGAAAC	ACGTTGGATGGTGAATTCAGTCCATGCAGG	ggacCAGGGTGAAACACTGTAAAC
BIEC590986	JPN_01	ACGTTGGATGGTTCTGCCAAGAAGAAATATG	ACGTTGGATGTTAGTGAATCATGCGTGGAG	gGATTAATGGAGGTGTTAAAAA
BIEC686800	JPN_01	ACGTTGGATGGCAATTTGTCCATCTCTGCTG	ACGTTGGATGGTCAGAGTATTTGTCCGTAG	ctgtTAGCAGTAATACTGTCACTA
BIEC266761	JPN_01	ACGTTGGATGAATCATCTCAGATGCGTCCC	ACGTTGGATGCCTGCAAAATCATGCCCTTG	gattCGTCCCTGGCCTAGTTTACCC
BIEC493879	JPN_01	ACGTTGGATGGGGACAATGACACTTCACAT	ACGTTGGATGGAGAGATGTGAGCCTGTTAC	ttACACTTCACATATTGTTATCACTT
BIEC86281	JPN_01	ACGTTGGATGAGAGCAACGTGAGAGAAGAG	ACGTTGGATGTGACCTTGGTCTTCTTGTCTG	ggggACGTGAGAGAAGAGCAGACAT
BIEC247284	JPN_01	ACGTTGGATGTCCCTCAAATGAACACAGC	ACGTTGGATGGGTGAACATATGTTTGAAGGC	gAGCATATGTGCAAAATCTATAAAAT
BIEC136821	JPN_01	ACGTTGGATGGGCACATTTCTGAGGTTTC	ACGTTGGATGGACTGTGCACAACAACCAG	attaTTGTGAAGGAAGAATGCAATTC
BIEC915102	JPN_01	ACGTTGGATGTCTTAACGAGCTGAAGAACTG	ACGTTGGATGTATTTGGTCAATCATCGTGG	AATTAATGTTTTCTCTATTAATCAAC
BIEC42118	JPN_01	ACGTTGGATGATGCTCATATATGACCTTG	ACGTTGGATGGTGCGTATGTGATGTGTCC	TCATAAAATTTCAATTAATGTCAAAAA
BIEC884767	JPN_01	ACGTTGGATGTCTTTGGGTGACCCGAATG	ACGTTGGATGCACCCACCATGAAACATAAG	gcggTGAATTTAATTCCTTACGATAAAA
BIEC562465	JPN_02	ACGTTGGATGAGATCCATCACTGTCTCTG	ACGTTGGATGGATGACTCATCTCTCTCTCC	CCTCCCTCTGCTGGCC
BIEC155175	JPN_02	ACGTTGGATGAATGTGAGAGGTCTTCCCC	ACGTTGGATGGTGTAGTTAATGCCAAACCG	GCCAAACCCGCTCAT
BIEC541693	JPN_02	ACGTTGGATGATTAACCCCTGGAATCAGCG	ACGTTGGATGACAGGAAGCAGCTCTGTGTG	TGCCAGGCACCTCTCC
BIEC554813	JPN_02	ACGTTGGATGTGAGCTCAATGGAAGAGG	ACGTTGGATGACATCAGTGTTCAGCTGTCC	CTGTCCAGGTGATCAG
BIEC798010	JPN_02	ACGTTGGATGACATGAGCCAAATGTCAACCC	ACGTTGGATGATTAGTCCACTTGGGGATAC	TTGGGGATACACTGGT
BIEC189021	JPN_02	ACGTTGGATGGAGAGCATGAGTGGAAATAGG	ACGTTGGATGGTGTGCAATAACCTTGACTTG	ecCTTGACTTGTCTCA
BIEC633979	JPN_02	ACGTTGGATGTCTACGATGCCAGAGTCAC	ACGTTGGATGGTTCGAATCCAAACATAAGCTG	TCCGTAGTTTGTCTTTCA
BIEC819385	JPN_02	ACGTTGGATGTCTCTCTCAGCTGTGTTTG	ACGTTGGATGATCTCCCTCCACAGTGTCTAC	ccCAGCTGAAGCATTCAA
BIEC324530	JPN_02	ACGTTGGATGTCTTTCAGTGTGCTCCTCAG	ACGTTGGATGGGGAATAGGGGAGAAGACTG	GAGAAGACTGAGCCAGAAG
BIEC204022	JPN_02	ACGTTGGATGCCTGAATCTATGCAAGGCAC	ACGTTGGATGTCCCAAGAAGACTGTTCTGC	TGCTAGATGACATCATCAIT
BIEC180122	JPN_02	ACGTTGGATGCAAAATGTTACGCCGCTCTG	ACGTTGGATGGATATTACAGAGGCAGCTGG	ttggtTGGCCGACCACGTTT
BIEC119158	JPN_02	ACGTTGGATGTCTTTCAGTGTGCTCCTCAG	ACGTTGGATGTGTACACTCCATCCAAACCAG	cccaTCTTGTCTTCTTGCTTA
BIEC349712	JPN_02	ACGTTGGATGGTGACACCAATTATGCAAAAC	ACGTTGGATGCTTACACTGCATGTATCTGTC	ATCTGTCAAGAATAGCAAAAG
BIEC531275	JPN_02	ACGTTGGATGCTTATCATCTATGGTGGTTG	ACGTTGGATGGGCCGACTCTTCTTAGCAA	ACTCTTCTTAGCAATGTTTAC
BIEC523483	JPN_02	ACGTTGGATGCTGGCTAAAGCAGAAAAATTC	ACGTTGGATGCCGTAGGTGGATGAGACTTG	gggtCTCATCTCATGTTGTTTC
BIEC884767	JPN_02	ACGTTGGATGTCCCAAGCTTTGGAAGATTG	ACGTTGGATGTTATCTGAGGCAGTTGGAAG	gGGAAGCAAAATTTGCTGAAGT
BIEC617070	JPN_02	ACGTTGGATGGATATTTCTCTCAAGCCG	ACGTTGGATGGCCGCTCTCTGATTCATCAA	gTGTGTGATTTCTGTGTTTCTC
BIEC571705	JPN_02	ACGTTGGATGGGGTCATCATTTTCAAGT	ACGTTGGATGTTATGTCAGAACGCTATCATC	gggtTAGATGATTACACTTTGCCC
BIEC431445	JPN_02	ACGTTGGATGGGGTCTGACCTTAGAAGTT	ACGTTGGATGGGATGTTTGAAGAGAAGAG	gggtAGAAAGTATTGCACAAAAAC
BIEC366938	JPN_02	ACGTTGGATGGAGTCCAATTTCCAAACAGAC	ACGTTGGATGCAAGAGAGTCTCTCTGAAG	ctatTCTCTGTAAGATTAACCTCT
BIEC183067	JPN_02	ACGTTGGATGCTTACCCACAGGAGCCAACAG	ACGTTGGATGTGGCACTCACATGCACACC	ggatGGGCGGCGCTCCGAATTCGCA
BIEC707898	JPN_02	ACGTTGGATGAAGAGGACTACGGGCTTTTG	ACGTTGGATGCACCTTGTGTAGCATCCGAGC	catcACAGACAATGTGATGATAATCC
BIEC420894	JPN_02	ACGTTGGATGTGGAGCTGACCAAGAAAGTG	ACGTTGGATGACACTTTTGAACAGCTACTG	ACGTGATGTGATTTTTCATATTTTAT
BIEC749477	JPN_02	ACGTTGGATGGGAAACCTGGGTTTGGCTC	ACGTTGGATGTGGGACCAAGACTTGACACAC	acccAGACTTGCACAAATTTTCTCTC
BIEC154171	JPN_02	ACGTTGGATGAGGCCATGCTTCTCTTTTG	ACGTTGGATGCATTTACCTACTGAAACTC	cccttAAACTCATAAATTTTGTGCTCTA
BIEC499860	JPN_02	ACGTTGGATGGACCAACATGGAGGAGATAG	ACGTTGGATGATGGGATCTGCACACAGAG	TGAATATTTTATATTTTAACTTGCATT

The sequence, "ACGTTGGATG", was added to all the PCR primers not to overlap molecular weights of the PCR primers and the extend primers. Small letters of the extend primers show artificial sequences to adjust a molecular weight of each allele for multiplex genotyping.

A) Panel: JPN_01 **B) Pane2: JPN_02** **C) Example of genotyping for BIEC696480 and BIEC581695**

Fig. 1. Mass spectrum profile of a horse genomic DNA sample analyzed using the SNP 53 JPN System with the MassARRAY® system for the (A) JPN_01, (B) JPN_02 panels and (C) example of SNP genotyping. Genotyped alleles were discriminated by differences of molecular weights of extended adenine (A), thymine (T), guanine (G) or cytosine (C). BIEC696480 and BIEC581695 have alleles "C" and "T".

culated from the allele frequencies. Three types of exclusion probabilities—PE01, PE02, and PE03—were calculated from the allele frequencies [13]. PE01 represented a one-parent exclusion case; i.e., the genotypes of a dam, her offspring, and a putative sire were compared. PE02 repre-

sented a missing-parent case; i.e., one of the parental genotypes was unavailable for testing. PE03 represented a both-parents exclusion case; i.e., a challenging that is an offspring falsely attributed to 2 parents was present, and all 3 genotypes were tested.

Table 2. Minor allele frequency (MAF), heterozygosity (H) and exclusion probability (PE) of the 53 SNPs

Local SNP ID	Chr.	Panel	MAF	H	PE01	PE02	PE03
BIEC42118	chr01	JPN_01	0.321	0.436	0.170	0.095	0.261
BIEC397614	chr02	JPN_01	0.253	0.378	0.153	0.071	0.240
BIEC672139	chr03	JPN_01	0.358	0.460	0.177	0.106	0.268
BIEC681989	chr03	JPN_01	0.411	0.484	0.183	0.117	0.276
BIEC811791	chr06	JPN_01	0.416	0.486	0.184	0.118	0.277
BIEC870244	chr08	JPN_01	0.421	0.488	0.184	0.119	0.277
BIEC871916	chr08	JPN_01	0.479	0.499	0.187	0.125	0.281
BIEC911841	chr09	JPN_01	0.279	0.402	0.161	0.081	0.249
BIEC915102	chr09	JPN_01	0.437	0.492	0.186	0.121	0.279
BIEC86281	chr10	JPN_01	0.479	0.499	0.187	0.125	0.281
BIEC136821	chr11	JPN_01	0.368	0.465	0.178	0.108	0.270
BIEC205561	chr14	JPN_01	0.242	0.367	0.150	0.067	0.236
BIEC247284	chr15	JPN_01	0.421	0.488	0.184	0.119	0.277
BIEC266761	chr16	JPN_01	0.426	0.489	0.185	0.120	0.278
BIEC322734	chr17	JPN_01	0.484	0.499	0.187	0.125	0.281
BIEC344848	chr19	JPN_01	0.426	0.489	0.185	0.120	0.278
BIEC493879	chr22	JPN_01	0.305	0.424	0.167	0.090	0.256
BIEC508410	chr23	JPN_01	0.321	0.436	0.170	0.095	0.261
BIEC532060	chr24	JPN_01	0.495	0.500	0.187	0.125	0.281
BIEC555737	chr25	JPN_01	0.374	0.468	0.179	0.110	0.271
BIEC581695	chr26	JPN_01	0.489	0.500	0.187	0.125	0.281
BIEC604433	chr27	JPN_01	0.158	0.266	0.115	0.035	0.192
BIEC590986	chr27	JPN_01	0.368	0.465	0.178	0.108	0.270
BIEC607490	chr28	JPN_01	0.463	0.497	0.187	0.124	0.280
BIEC633181	chr29	JPN_01	0.311	0.429	0.168	0.092	0.258
BIEC686800	chr30	JPN_01	0.232	0.356	0.146	0.063	0.232
BIEC696480	chr31	JPN_01	0.468	0.498	0.187	0.124	0.281
Average			0.378	0.454	0.175	0.105	0.266

Local SNP ID	Chr.	Panel	MAF	H	PE01	PE02	PE03
BIEC366938	chr02	JPN_02	0.342	0.450	0.174	0.101	0.265
BIEC420894	chr02	JPN_02	0.289	0.411	0.163	0.084	0.252
BIEC707898	chr04	JPN_02	0.384	0.473	0.181	0.112	0.273
BIEC798010	chr06	JPN_02	0.426	0.489	0.185	0.120	0.278
BIEC819385	chr07	JPN_02	0.363	0.462	0.178	0.107	0.269
BIEC884767	chr08	JPN_02	0.395	0.478	0.182	0.114	0.274
BIEC119158	chr10	JPN_02	0.421	0.488	0.184	0.119	0.277
BIEC154171	chr12	JPN_02	0.211	0.333	0.139	0.055	0.222
BIEC155175	chr12	JPN_02	0.258	0.383	0.155	0.073	0.242
BIEC180122	chr13	JPN_02	0.379	0.471	0.180	0.111	0.272
BIEC183067	chr14	JPN_02	0.458	0.496	0.187	0.123	0.280
BIEC189021	chr14	JPN_02	0.463	0.497	0.187	0.124	0.280
BIEC204022	chr14	JPN_02	0.232	0.356	0.146	0.063	0.232
BIEC324530	chr18	JPN_02	0.347	0.453	0.175	0.103	0.266
BIEC349712*	chr19	JPN_02	0.463	0.529	0.230	0.140	0.347
BIEC431445	chr20	JPN_02	0.258	0.383	0.155	0.073	0.242
BIEC479477	chr22	JPN_02	0.226	0.350	0.144	0.061	0.229
BIEC499860	chr22	JPN_02	0.447	0.494	0.186	0.122	0.279
BIEC541693	chr24	JPN_02	0.253	0.378	0.153	0.071	0.240
BIEC523483	chr24	JPN_02	0.326	0.439	0.171	0.097	0.262
BIEC531275	chr24	JPN_02	0.268	0.392	0.158	0.077	0.245
BIEC554813	chr25	JPN_02	0.374	0.468	0.179	0.110	0.271
BIEC562465	chr26	JPN_02	0.300	0.420	0.166	0.088	0.255
BIEC571705	chr26	JPN_02	0.500	0.500	0.188	0.125	0.281
BIEC617070	chr28	JPN_02	0.284	0.407	0.162	0.083	0.250
BIEC633979	chr29	JPN_02	0.400	0.480	0.182	0.115	0.275
Average			0.349	0.442	0.173	0.099	0.264
JPN_01 (27 SNPs)					0.99442	0.94992	0.99976
JPN_02 (26 SNPs)					0.99281	0.93391	0.99966
Total					0.99996	0.99669	0.99999

* MAF in BIEC349712 shows the total allele frequency of the second and third alleles, because of having 3 alleles.

Table 3. Allele number, heterozygosity (H), exclusion probability (PE) of the STRs

Locus	Chr.	Panel	Allele	H	PE01	PE02	PE03
VHL20	chr30	kakoi's set	4	0.733	0.484	0.310	0.659
HTG4	chr09	kakoi's set	4	0.534	0.249	0.144	0.377
AHT4	chr24	kakoi's set	4	0.713	0.459	0.288	0.634
HMS7	chr01	kakoi's set	5	0.777	0.560	0.381	0.740
CA425	chr28	kakoi's set	6	0.601	0.366	0.199	0.546
TKY28	chr06	kakoi's set	6	0.695	0.454	0.284	0.638
HTG10	chr21	kakoi's set	6	0.810	0.623	0.446	0.802
AHT5	chr08	kakoi's set	5	0.698	0.458	0.283	0.641
HMS3	chr09	kakoi's set	5	0.621	0.397	0.221	0.589
TKY19	chr18	kakoi's set	4	0.708	0.460	0.286	0.638
ASB017	chr02	kakoi's set	5	0.744	0.507	0.330	0.686
ASB023	chr03	kakoi's set	6	0.770	0.555	0.376	0.739
HMS6	chr04	kakoi's set	4	0.613	0.332	0.193	0.485
LEX033	chr04	kakoi's set	5	0.702	0.451	0.280	0.628
ASB002	chr15	kakoi's set	8	0.810	0.626	0.450	0.806
Average			5.1	0.702	0.465	0.298	0.641
Total PE					0.99994	0.99560	>0.99999
TKY279	chr16	15 TKY	6	0.751	0.526	0.348	0.710
TKY287	chr17	15 TKY	4	0.644	0.404	0.231	0.586
TKY294	chr27	15 TKY	5	0.729	0.483	0.308	0.659
TKY297	chr01	15 TKY	5	0.742	0.506	0.331	0.687
TKY301	chr23	15 TKY	7	0.751	0.524	0.347	0.706
TKY312	chr06	15 TKY	7	0.702	0.468	0.292	0.655
TKY321	chr20	15 TKY	7	0.806	0.621	0.444	0.804
TKY325	chr29	15 TKY	7	0.738	0.513	0.337	0.701
TKY333	chr28	15 TKY	6	0.783	0.581	0.402	0.766
TKY337	chr04	15 TKY	4	0.720	0.468	0.296	0.642
TKY341	chr16	15 TKY	6	0.650	0.416	0.243	0.604
TKY343	chr11	15 TKY	7	0.732	0.497	0.322	0.680
TKY344	chr05	15 TKY	6	0.774	0.565	0.387	0.751
TKY374	chr01	15 TKY	7	0.781	0.579	0.401	0.765
TKY394	chr24	15 TKY	6	0.805	0.613	0.435	0.792
Average			6.0	0.741	0.518	0.342	0.701
Total PE					0.99998	0.99823	>0.99999

RESULTS

Construction of multiplex assays for genotyping SNPs: First, 60 polymorphic SNPs were selected from 120 SNPs, which had been randomly selected from the entire set of horse chromosomes, by genotyping 8 thoroughbreds. Two multiplex sets for the iPLEX assay were designed using the MassARRAY[®] assay design software; the 60 SNPs were separated into 2 panels—pre-JPN_01 (31 SNPs) and pre-JPN_02 (29 SNPs). After analysis with MassARRAY[®] software using 16 thoroughbreds the iPLEX assay was optimized by excluding 7 SNPs from the multiplex system because of unstable amplification and/or primer extension. Finally, 2 stable multiplex sets for the iPLEX assay were constructed for genotyping the remaining 53 SNPs; these SNPs were separated into 2 panels—JPN_01 (27 SNPs) and JPN_02 (26 SNPs). The primers designed for PCR and extension are shown in Table 1.

Figure 1 shows the results of amplification and primer extension of genomic DNA with JPN_01 (Fig. 1A) and

JPN_02 (Fig. 1B) SNPs. Figure 1C shows examples for SNP-genotyping discriminated by differences of molecular weights of cytosine (C) and thymine (T). All the SNPs of the multiplex sets were well amplified and could be genotyped. Some nonspecific artifacts might have been amplified along with the products specific for each SNP. These artifacts, however, did not affect the SNP-specific products. These results indicate that the SNP panels forming the SNP 53 JPN System, would be useful for parentage testing of thoroughbreds.

Random samples from 95 thoroughbreds, which included a pseudo parent and sib combination from the 2006 ISAG horse comparison test, were used to genotype the 53 SNPs. Minor allele frequency (MAF), heterozygosity (H), and exclusion probability (PE01, PE02, and PE03) [13] of the population are shown in Table 2. The SNPs, in which 1 SNP (BIEC349712) had 3 alleles, were found to be highly informative markers as indicated by the 26–53% heterozygosity. The exclusion probability (PE01) in combination with the system for thoroughbreds was 99.996% (Table 2),

Table 4. The abilities of parentage exclusion between SNPs and STRs

SNP 53 JPN system
JPN_01 (27 SNPs)

Sample	#42118	#397614	#672139	#681989*	#811791	#870244	#871916	#911841	#915102*	#86281	#136821	#205561	#247284	#266761
HCT06_01	TT	CC	CT	CC	TT	CC	AT	GC	TT	TT	AA	GG	AG	CT
HCT06_05	TT	CT	CT	TT	TT	GC	AT	GC	CC	TT	AA	GG	AG	CC

JPN_02 (26 SNPs)

Sample	#366938*	#420894	#707898	#798010*	#819385	#884767	#119158	#154171	#155175	#180122*	#183067	#189021	#204022
HCT06_01	AA	TT	TT	AA	CC	GT	AA	TT	CT	GG	AA	GT	AA
HCT06_05	GG	CT	CT	GG	CA	TT	AA	TT	CT	AA	AG	TT	AA

Sample	#324530	#349712	#431445	#479477	#499860*	#541693	#523483*	#531275	#554813	#562465	#571705	#617070	#633979
HCT06_01	TC	AT	TC	CC	GG	CC	TT	CT	GG	CC	AG	TT	CC
HCT06_05	TC	AT	TC	CC	TT	CT	CC	CT	GG	CC	AA	TC	CT

Microsatellite 15 TKY system

Sample	TKY344	TKY279*	TKY343	TKY321*	TKY287	TKY312*	TKY301	TKY337*	TKY374*	TKY297	TKY333*	TKY341*	TKY325	TKY294	TKY394
HCT06_01	IK	MM	MM	IQ	NN	II	MN	LM	JK	LO	KT	KK	OP	PP	JO
HCT06_05	IQ	JN	MM	LS	NR	MN	NO	OO	LO	MO	SS	MQ	JP	MP	JL

Kakoi's panel including the minimum 9 markers

Sample	VHL20	HTG4	AHT4	HMS7*	CA425	TKY28*	HTG10	AHT5	HMS3	TKY19	ASB17	ASB23	HMS6*	LEX33*	ASB2
HCT06_01	MN	MP	JK	NO	LN	1, 4	IO	JM	II	LP	GO	JS	MM	LO	KQ
HCT06_05	LM	KM	JO	JM	IN	2, 5	II	JN	II	LL	GG	LS	PP	MM	KM

*Asterisk shows the markers that excluded the pseudo parent and sib combination.

which was comparable to that for systems that use STRs, such as the Microsatellite 15 TKY System. The discrimination power of a single SNP was less than that of an STR, but a combination of several SNPs resulted in a high discriminating power. Those results were used to evaluate the Hardy-Weinberg (HW) equilibrium. Out of the 53 SNPs, 50 loci conformed to the HW equilibrium ($P > 0.05$); however, 3 loci, namely, BIEC811791, BIEC590986, and BIEC633979, showed deviation from the HW equilibrium ($P < 0.05$). The PE01 of the 50 SNPs that conformed to the HW equilibrium was 99.994%, indicating that only these 50 SNPs were adequate for parentage testing of thoroughbreds.

Genotyping of STRs: Two parentage testing systems—the STRs reported by Kakoi (including the minimum 9 markers recommended by the ISAG) and the Microsatellite 15 TKY System—were used to genotype the same 95 thoroughbreds used for the SNP analysis. All the STRs of the 2 systems were well amplified and could be genotyped since they had been routinely utilized for horse parentage testing in our laboratory. In accordance with the literature [15, 33], many STRs showed high heterozygosity and multiple alleles (Table 3). In this study, the exclusion probability (PE01) for the STRs reported by Kakoi and for the Microsatellite 15 TKY System for thoroughbreds was 99.994% and 99.998%, respectively.

Parentage testing cases: The 3 systems for parentage testing—the SNP 53 JPN System for SNP genotyping, the

STRs reported by Kakoi and the Microsatellite 15 TKY System for STR genotyping—were used to evaluate the exclusion of an actual parentage testing case, which was a pseudo parent and sib combination found using the 2006 ISAG horse comparison test. In the SNP 53 JPN System, 8 of the 53 SNPs excluded the parent and sib combination as a pseudo combination; 4 STRs of the STRs reported by Kakoi and 7 STRs in the Microsatellite 15 TKY System also excluded it as a pseudo combination (Table 4).

DISCUSSION

The objective of the present study was to construct a parentage testing system using SNPs for parentage testing of thoroughbreds and to evaluate its utility for routine parentage testing by comparing it to the present system based on STRs [15, 33]. The ISBC requires that an exclusion probability (PE01) above 99.95% be maintained when switching from the blood typing technique [1–4, 8, 16, 24, 25, 31, 32] to STR typing for routine parentage testing of thoroughbreds. Thus, we aimed to construct an SNP-based parentage testing system with PE01 above 99.95%. In this study, we employed a multiplex assay using single-nucleotide primer extension reactions for SNP genotyping because a multiplex assay is convenient and requires only small amounts of DNA; moreover, the primer extension reaction is simple and cost-effective. The SNP 53 JPN System fulfilled the total

PE value requirements of the ISBC for parentage testing, and it could be used to clearly genotype 53 SNPs by using only 2 multiplex sets, JPN_01 (27 SNPs) and JPN_02 (26 SNPs). Additionally, this system could exclude the pseudo combination found with the ISAG horse comparison test, which uses multiple markers, as well as the current parentage testing systems, which use STRs: this suggests that our system is suitable for thoroughbred parentage testing.

In this study, 3 SNPs showed deviation from the HW equilibrium for thoroughbreds. Although the causes for this deviation are unknown, it is known that multiple genomic regions are evolutionarily biased in thoroughbreds [9]. The 3 SNPs may be located in these biased genomic regions. The exclusion probability (PE01) [13] of the remaining 50 SNPs was also above 99.95% as required by the ISBC: this indicates that our system was adequate for parentage testing of thoroughbreds.

Although several different assays are available for SNP genotyping, such as the TaqMan[®] assay [12] and DHPLC [20], the iPLEX assay along with the MassARRAY[®] [14] was suitable for routine parentage testing because it enabled simultaneous genotyping of multiple SNPs. The current panel of 53 SNPs was found to be sufficient for parentage testing of thoroughbreds, and genotyping performed using the iPLEX assay with MassARRAY[®] was as cost- and time-effective as that by the STR assay using automated sequencers. The exclusion probabilities computed for this panel of SNPs were similar to those reported using STRs for thoroughbreds [15, 33], indicating that the discrimination power of about 50 SNPs is equivalent to that of about 15 STRs.

One of the main proposed benefits of SNP-based systems is the potential for genetic diagnosis, such as the determination of coat color [5–7, 10, 11, 21, 27, 30] and diseases in horses. It is expected that the genetic diagnosis of coat color and diseases would help the actual coat color information and clinical symptom gained by observation, since coat color information is particularly now available for horse identification and registry. Although the SNP 53 JPN System includes 53 SNPs in its panels, it is possible to include additional SNPs in the panels because the current iPLEX assay technology can accommodate over 40 SNPs in 1 multiplex assay (panel). Therefore, this is a flexible technology that can be improved to serve further demands, for example, for the genotyping of factors such as coat color genotyping.

This is the first report on horse parentage testing with multiplex assays using many SNPs: it is possible that our system, the SNP 53 JPN System, will prove to be a basic tool for improving and developing the current parentage testing system, which uses STRs, for horses.

ACKNOWLEDGMENT(S). We thank Dr. M. Kurosawa from the Laboratory of Racing Chemistry and Dr. N. Ishida of the Japan Racing Association, for valuable discussions. This study was supported by a grant-in-aid from the Japan Racing Association (2008).

REFERENCES

1. Bell, K., Patterson, S. and Pollitt, C. C. 1984. The plasma protease inhibitor system (Pi) of Standardbred horses. *Anim. Blood Groups Biochem. Genet.* **15**: 191–206.
2. Bell, K., Pollitt, C. C. and Patterson, S. D. 1988. Subdivision of equine Tf into H₁ and H₂. *Anim. Genet.* **19**: 177–183.
3. Braend, M. 1967. Variation of horse prealbumins in acidic starch gels. *Acta Vet. Scand.* **8**: 193–194.
4. Braend, M. and Johansen, K. E. 1983. Haemoglobin types in Norwegian horses. *Anim. Blood Groups Biochem. Genet.* **14**: 305–307.
5. Brooks, S. A., Terry, R. B. and Bailey, E. 2002. A PCR-RFLP for KIT associated with tobiano spotting pattern in horses. *Anim. Genet.* **33**: 301–303.
6. Brooks, S. A. and Bailey, E. 2005. Exon skipping in the KIT gene causes a Sabino spotting pattern in horses. *Mamm. Genome* **16**: 893–902.
7. Cook, D., Brooks, S., Bellone, R. and Bailey, E. 2008. Missense mutation in exon 2 of SLC36A1 responsible for champagne dilution in horses. *PLoS Genet.* **4**: e1000195.
8. Gahne, B. and Juneja, R. K. 1978. Polymorphic post-albumin of cattle and horse plasma identified as vitamin D binding protein (Gc protein). *Anim. Blood Groups Biochem. Genet.* **9**: 37–40.
9. Gu, J., Orr, N., Park, S. D., Katz, L. M., Sulimova, G., MacHugh, D. E. and Hill, E. W. 2009. A genome scan for positive selection in thoroughbred horses. *PLoS ONE* **4**: e5767.
10. Haase, B., Brooks, S. A., Schlumbaum, A., Azor, P. J., Bailey, E., Alaeddine, F., Mevissen, M., Burger, D., Poncet, P. A., Rieder, S. and Leeb, T. 2007. Allelic heterogeneity at the equine KIT locus in dominant white (W) horses. *PLoS Genet.* **3**: e195.
11. Haase, B., Brooks, S. A., Tozaki, T., Burger, D., Poncet, P. A., Rieder, S., Hasegawa, T., Penedo, C. and Leeb, T. 2009. Seven novel KIT mutations in horses with white coat colour phenotypes. *Anim. Genet.* **40**: 623–629.
12. Hui, L., DelMonte, T. and Ranade, K. 2008. Genotyping using the TaqMan assay. *Curr. Protoc. Hum. Genet. Chapter 2*: Unit 2.
13. Jamieson, A. and Taylor, S. C. 1997. Comparisons of three probability formulae for parentage exclusion. *Anim. Genet.* **28**: 397–400.
14. Jurinke, C., van den Boom, D., Cantor, C. R. and Köster, H. 2002. The use of MassARRAY technology for high throughput genotyping. *Adv. Biochem. Eng. Biotechnol.* **77**: 57–74.
15. Kakoi, H., Nagata, S. and Kurosawa, M. 2001. DNA typing with 17 microsatellites for parentage verification of racehorses in Japan. *Anim. Sci. J.* **72**: 453–460.
16. Kaminski, M. and Gajos, E. 1964. Comparative examination of carboxylic esterases in sera of horse, donkey and their hybrids. *Nature* **201**: 716–718.
17. Lee, H. Y., Park, M. J., Yoo, J. E., Chung, U., Han, G. R. and Shin, K. J. 2005. Selection of twenty-four highly informative SNP markers for human identification and paternity analysis in Koreans. *Forensic Sci. Int.* **148**: 107–112.
18. Lessig, R., Zoledziewska, M., Fahr, K., Edelmann, J., Kos-trzewa, M., Dobosz, T. and Kleemann, W. J. 2005. Y-SNP-genotyping—a new approach in forensic analysis. *Forensic Sci. Int.* **154**: 128–136.
19. Litt, M. and Luty, J. A. 1989. A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. *Am. J. Hum. Genet.* **44**:

- 397–401.
20. Liu, W., Smith, D. I., Rechtzigel, K. J., Thibodeau, S. N. and James, C. D. 1998. Denaturing high performance liquid chromatography (DHPLC) used in the detection of germline and somatic mutations. *Nucleic Acids Res.* **26**: 1396–1400.
21. Mariat, D., Taourit, S. and Guérin, G. 2003. A mutation in the MATP gene causes the cream coat colour in the horse. *Genet. Sel. Evol.* **35**: 119–133.
22. Mein, C. A., Barratt, B. J., Dunn, M. G., Siegmund, T., Smith, A. N., Esposito, L., Nutland, S., Stevens, H. E., Wilson, A. J., Phillips, M. S., Jarvis, N., Law, S., de Arruda, M. and Todd, J. A. 2000. Evaluation of single nucleotide polymorphism typing with invader on PCR amplicons and its automation. *Genome Res.* **10**: 330–343.
23. McKay, S. D., Schnabel, R. D., Murdoch, B. M., Matukumalli, L. K., Aerts, J., Coppeters, W., Crews, D., Dias Neto, E., Gill, C. A., Gao, C., Mannen, H., Wang, Z., Van Tassell, C. P., Williams, J. L., Taylor, J. F. and Moore, S. S. 2008. An assessment of population structure in eight breeds of cattle using a whole genome SNP panel. *BMC Genet.* **9**: 37.
24. Nogaj, A., Duniec, M. J., Slota, E. and Duniec, M. 1997. Three newly detected alloantigens in the U blood group system of horses. *Anim. Genet.* **28**: 313–314.
25. Op't Hof, J. and Osterhoff, D. R. 1973. Isoenzyme polymorphism of 6-phosphogluconate dehydrogenase (EC 1.1.1.44) in the family Equide. *Anim. Blood Groups Biochem. Genet.* **4**: 111–113.
26. Petkov, P. M., Ding, Y., Cassell, M. A., Zhang, W., Wagner, G., Sargent, E. E., Asquith, S., Crew, V., Johnson, K. A., Robinson, P., Scott, V. E. and Wiles, M. V. 2004. An efficient SNP system for mouse genome scanning and elucidating strain relationships. *Genome Res.* **14**: 1806–1811.
27. Rieder, S., Taourit, S., Mariat, D., Langlois, B. and Guérin, G. 2001. Mutation in the agouti (ASIP), the extension (MC1R), and the brown (TYRP1) loci and their association to coat color phenotypes in horses (*Equus caballus*). *Mamm. Genome* **12**: 450–455.
28. Rohrer, G. A., Freking, B. A. and Nonneman, D. 2007. Single nucleotide polymorphisms for pig identification and parentage exclusion. *Anim. Genet.* **38**: 253–258.
29. Salisbury, B. A., Pungliya, M., Choi, J. Y., Jiang, R., Sun, X. J. and Stephens, J. C. 2003. SNP and haplotype variation in the human genome. *Mutat. Res.* **15**: 53–61.
30. Santschi, E. M., Purdy, A. K., Valberg, S. J., Vrotsos, P. D., Kaese, H. and Mickelson, J. R. 1998. Endothelin receptor B polymorphism associated with lethal white foal syndrome in horses. *Mamm. Genome* **9**: 306–309.
31. Stormont, C. and Suzuki, Y. 1963. Genetic control of albumin phenotypes in horses. *Proc. Soc. Exp. Biol. Med.* **114**: 673–675.
32. Stormont, C. and Suzuki, Y. 1964. Genetic system of blood groups in horses. *Genetics* **50**: 915–929.
33. Tozaki, T., Kakoi, H., Mashima, S., Hirota, K., Hasegawa, T., Ishida, N., Miura, N., Choi-Miura, N. H. and Tomita, M. 2001. Population study and validation of paternity testing for Thoroughbred horses by 15 microsatellite loci. *J. Vet. Med. Sci.* **63**: 1191–1197.
34. Tozaki, T., Mashima, S., Hirota, K., Miura, N., Choi-Miura, N. H. and Tomita, M. 2001. Characterization of equine microsatellites and microsatellite-linked repetitive elements (eMLREs) by efficient cloning and genotyping methods. *DNA Res.* **28**: 33–45.
35. Weber, J. L. and May, P. E. 1989. Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *Am. J. Hum. Genet.* **44**: 388–396.
36. Werner, F. A., Durstewitz, G., Habermann, F. A., Thaller, G., Krämer, W., Kollers, S., Buitkamp, J., Georges, M., Brem, G., Mosner, J. and Fries, R. 2004. Detection and characterization of SNPs useful for identity control and parentage testing in major European dairy breeds. *Anim. Genet.* **35**: 44–49.